# Candidate Elimination Algorithm

→ It is used to find consistent hypothesis

| | Example | Sky | Air Temp | Humidity | Wind | Water | Forecast |
|---|---|---|---|---|---|---|---|
| +ve | 1 | Sunny | Warm | Normal | Strong | Warm | Same |
| +ve | 2 | Sunny | Warm | High | Strong | Warm | Same |
| -ve | 3 | Rainy | Cold | High | Strong | Warm | Change |
| +ve | 4 | Sunny | Warm | High | Strong | Cool | Change |

Most specific hypothesis $S_0 = \{\phi, \phi, \phi \cdots >$

a generic hypothesis $G_0 = \{?, ?, ?, ?, ?, ?\}$

For positive samples, we need to move from Specific to general.

For negative samples, we need to move from general to specific

$S_0 = \{\phi, \phi, \phi, \phi, \phi, \phi\}$

$G_0 = \{?, ?, ?, ?, ?, ?\}$

1) Check whether $x_1$ is +ve (specific to general)

As we dont have no predefined hypothesis

$S_1 = \{$'Sunny', 'warm', 'normal', 'strong', 'warm', 'same'$\}$

$G_1 = \{?, ?, ?, ?, ?, ?\}$

2) +ve – Specific to general

Example 2

$S_2 = \{$ 'Sunny', 'warm', ?, 'strong', 'warm', Same $\}$ with

$G_2 : \{ ?, ?, ?, ?, ?, ?, ? \}$

$G_2$ (general hypothesis will remain same un

and unless we get –ve example

3) $S_3 = \{$ 'Sunny', 'warm', ?, 'strong', 'warm',

As ex3 is –ve, change generic to specifi

so $S_3$ will not be changed. It is taken

from $S_2$.

Compare ~~$e_{x_3}$~~ with ex3

$E_{x3} \rightarrow \{$ Rainy, Cold, High, Strong, Warm, Change $\}$

If any change in ex3 data you need to represent value which is in $S_3$.

$E_{x3} = \{$ 'Sunny', 'warm', High, 'strong', 'warm', 'Same' $\}$

$G_3 = \{ <$'Sunny', ? ? ? ? ? $>, <$ ?, 'warm', ?, ?, ?, ? $>$

$<$ ?, ?, ?, ?, ?, 'Same' $>$

(ve) → Specific to general

$S_4$ => $S_3$ compare with Ex 4

$S_4$ = { 'sunny', 'warm', ?, 'strong', ?, ? >

if value matches same infn is written

" " doesn't match then replace with ?

$G_4$ = As it is +ve value same $G_3$ data should be written but same is not present in $S_4$ we need to remove 'same' attribute from $G_3$.

$G_4$ = { '<sunny', ?, ?, ?, ?, ? > < ?, 'warm', ?, ?, ? ?>}

$S_4$ and $S_4$ are final hypothesis

# Candidate Elimination Algorithm

## Example Problem

| Size | Color | Shape | class / Label |
|---|---|---|---|
| Big | Red | Circle | No |
| small | Red | Triangle | No |
| small | Red | Circle | Yes |
| Big | Blue | Circle | No |
| small | Blue | circle | Yes |

Ans: S:G : (small, ?, Circle)

$S_0 = (\emptyset, \emptyset, \emptyset)$

$G_0 = (?, ?, ?)$

$x_1 \rightarrow$ -ve example — No change

$S_1 = (\emptyset, \emptyset, \emptyset)$

$x_1 = (big, red, circle)$

As it big we need to check another minimal specializations.

There is another value "small".

$G_1 = \{ \langle small, ?, ? \rangle, \langle ?, Blue, ? \rangle \langle ?, ?, Triangle \rangle \}$

Why Blue: As there is no other possible value, we will replace Red with Blue

Why Triangle: As there is Yes and No for "Circle", we can write minimal

possible way ie triangle

$\therefore G_1 = \{ \langle small, ?, ? \rangle \langle ?, Blue, ? \rangle$

$\langle ?, ?, Triangle \rangle \}$

Ex 2 $x_2$ - (small, red, triangle) → Negative

$S_2 = (\phi, \phi, \phi)$ — As small and Triangle are not

$G_2 = \{ (?, Blue, ?) \}$

Compare with $G_1$ with Ex 2 to get $G_2$.

Ex 3 $x_3 =$ (small, red, circle)

(+ve example.

$S_3 =$ (small, red, circle)

$G_3 = \emptyset$ → when you compare $G_2$ with $S_3$ not attributes are matched.

Candidate elimination cannot allow G to be empty, because G must represent all maximally general hypothesis consistent with data so far.

So, we must reconstruct G.

ever G is empty, candidate elimination algorithm will not allow. It generates new hypothesis by excluding previous -ve examples.

-ve examples

$x_1$ — (Big, red, circle)

$x_2$ - (Small, red, triangle)

So, Size ≠ Big

Shape ≠ Triangle

Color must be Red (from Positives)

$S_3$ = (small, red, circle)

1) Generalize color

(small, ?, circle)

Check with negative examples

$x_1$ — (Big, red, circle)

$x_2$ - (small, red, triangle)

Size, shape — mismatch so it is valid.

2) Generalize size

(?, Red, circle)

$x_1$ — Big, red, circle)

# CANDIDATE ELIMINATION ALGORITHM

| ncome | Credit History | Employment | Approve |
|---|---|---|---|
| High | Good | Permanent | Yes |
| Medium | Good | Permanent | Yes |
| Low | Bad | Temporary | NO |
| Medium | Bad | Temporary | NO |

$S_0 = (\phi, \phi, \phi)$

$G_0 = (?, ?, ?)$

$S_1 = (High, Good, Permanent)$ +ve example

$G_1 = (?, ?, ?)$

$S_2 = (?, Good, Permanent)$ $S_2$ Compare with $I_2$

(+ve example)

$G_2 = (?, ?, ?)$

$S_3 = (?, Good, Permanent)$

As it is -ve example $\boxed{S_3 = S_2}$

$G_3 = (<Medium, ?, ?> <?, Good, ?> <?, ?, Permanent>)$

As Medium is not present in $S_3$, we can eliminate.

$\therefore G_3 = (<?, Good, ?> <?, ?, Permanent>)$

$x_4$ - -ve example ( Medium, Bad,

$S_4$ — ( ?, Good, Permanent )

$S_4$ is not changed $\boxed{S_4 = g_2}$

$G_{34}$ = ( < ?, Good, ? >, < ?, ? Permanent > )

$G_4$ = remains unchanged = $G_3$

$S_4$ = ( ?, Good, Permanent )

$G_4$ = ( < ?, Good, ? > < ?, ? Permanent > )

# Candidate Elimination Algorithm

## Example Problem

| Example | Citations | Size | In library | Price | Editions | Buy |
|---|---|---|---|---|---|---|
| 1 | Some | Small | No | Affordable | One | No |
| 2 | Many | Big | No | Expensive | Many | Yes |
| 3 | Many | Medium | No | Expensive | Few | Yes |
| 4 | Many | Small | No | Affordable | Many | Yes |

$S_4$ = (Many, ?, No, ?, ?)  $G_4$ = (Many, ?, ?, ?, ?)

$S_0$ = (ϕ, ϕ, ϕ, ϕ, ϕ)

$G_0$ = (?, ?, ?, ?, ?)

$e_1$ - 1 : -ve example

$S_1$ = (ϕ, ϕ, ϕ, ϕ, ϕ)

As ϕ doesn't match with some, small, No, ...

$G_1$ = As we know general hypothesis '?' matches with any value - so it is +ve classification. But $e_1$-1 is -ve result. So find out possible hypothesis by replacing one value at single time with opposite value.

∴ $G_1$ = (Many, ?, ?, ?, ?) (?, Big, ?, ?, ?) (?, Medium, ?, ?, ?)

(? ? ? Exp, ?) (? ? ? ? Many) (? ? ? ? Few)

$S_1$ = (ϕ, ϕ, ϕ, ϕ, ϕ)

Ex-2 — +ve example

$S_2$ : (Many, Big, No, Expensive, Many)

$G_2$ : (Many, ?, ? ? ?) (? Big ? ? ?) (? ? ? Exp ?)

(? ? ? ? Many)

(Few, medium doesn't match)

Ex-3 — +ve example

$G_3$ = (Many, ? ? ? ?) (? ? ? Exp ?)

(Few, Big doesn't match so removed)

$S_3$ = (Many, ?, No, Exp, ?)

Ex-4 — +ve example

$G_4$ = (Many, ? ? ? ?)

(Exp doesn't match)

$S_4$ = (Many, ?, No, ?, ?)

∴ $G_4$ = (Many, ?, ?, ?, ?)

$S_4$ = (Many, ?, No, ?, ?)

# Candidate Elimination Algorithm

| Income | Credit | Employment | Approve |
|---|---|---|---|
| High | Good | Permanent | Yes |
| Medium | Good | Permanent | Yes |
| Low | Bad | Temporary | No |
| Medium | Bad | Temporary | NO |

$S_0 = (\phi, \phi, \phi)$

$G_0 = (?\ ?\ ?)$

Ex-1 $\longrightarrow$ +ve

$S_1 = (High, Good, Permanent)$

$G_1 = (?\ ?\ ?)$

Ex-2 $\rightarrow$ +ve

$S_2 = (?, Good, Permanent)$

$G_2 = (?, ?, ?)$

Ex-3 . . . -ve example

$G_3 \Rightarrow$ All ? matches with values. So +ve Classification but result -ve. So not consistent. find possible hypothesis

$G_3 = (High, ?, ?)\ (Medium, ?, ?)\ (?, Good, ?)\ (?\ ?\ Permanent)$

$S_3 = (?, Good, Permanent$

Ex-4 - ve example.

$G_3$ = (High, ? ?) (Medium, ? ?) (? Good ? ?) (? ? Permanent)

$G_4$ => ① Medium doesn't match with high
-ve classification → -ve o/p ✓

② Medium matches with medium
↳ +ve → -ve o/p X

③ Good doesn't match - ve class → -ve o/p ✓

④ doesn't match - ✓

$G_4$ = (~~Medium~~ (High, ?, ?) (? Good ?) (? ? Permanent)

$S_4$ = (?, Good, Permanent).

∴ $S_4$ = (?, Good, Permanent)

$G_4$ = { (High, ?, ?) (?, Good, ?) (?, ?, Permanent) }

Eliminate.

# Candidate Elimination Algorithm

| Example | Shape | Size | Color | Surface | Thickness | Target |
|---|---|---|---|---|---|---|
| 1 | Circular | large | light | smooth | thick | Yes |
| 2 | Circular | large | light | irregular | thick | Yes |
| 3 | Oval | large | dark | smooth | thin | NO |
| 4 | Oval | large | light | irregular | thick | Yes |

$S_0 = (\emptyset, \emptyset, \emptyset, \emptyset, \emptyset)$

$G_0 = (?, ?, ?, ?, ?)$

Ex-1 — +ve example

$S_1 = (\text{circular}, \text{large}, \text{light}, \text{smooth}, \text{thick})$

$G_1$ — ? matches with all — +ve classification +ve o/p

$G_1 = (?, ?, ?, ?, ?)$

Ex-2 +ve example

$S_2 = (\text{circular}, \text{large}, \text{light}, ?, \text{thick})$

$G_2 = (?, ?, ?, ?, ?)$

Ex-3 — -ve example

Ex-3 (oval, large, dark, smooth, thin)

$G_3$ = All ? matches with all — +ve classification but -ve o/p.

Not consistent. We need to find all possible hypothesis

| | | | |
|---|---|---|---|
| New | No | | |
| Returning | No | Positive | Yes |

$G_3$ = ( Circular, ? ? ? ? ) ( ? ? light ? ? )
( ? ? ? ? thick )

eliminate irregular

$S_3$ = (Circular, large, light, ?, Thick)

10-12

Ex-4 — +ve example

$S_4$ = ( ?, large, light, ?, thick )

$G_4$ = ① ( Circubr, ? ? ? ? ) — no match remove

② ( ? ? light ? ? ) — match ✓

~~③ ( ? ? ? irregular ? ) — match~~ ✓

④ ( ? ? ? ? thick ) — match ✓

~~$G_4$ = ( Circular, ? ? ? ? ) ( ? ? light ? ? ) ( ? ? ? irregular ? )~~

(

$G_4$ = ( ? ? light ? ? ) ( ? ? ? irregular ? ) ( ? ? ? ? thick )

∴ $S_4$ = ( ?, large, light, ?, Thick )

$G_4$ = { ( ? ? light ? ? ) ( ? ? ? irregular ? ) ( ? ? ? ? thick )
eliminate

$G_4$ = ( ? ? light ? ? ) ( ? ? ? ? thick )

# Candidate Elimination Algorithm

| Size | Color | Shape | Class |
|---|---|---|---|
| Big | Red | Circle | NO - 1 |
| small | Red | Triangle | NO - 2 |
| Small | Red | Circle | Yes - 3 |
| Big | Blue | Circle | NO - 4 |
| small | Blue | Circle | Yes - 5 |

$S_0 = (\phi, \phi, \phi)$

$G_0 = (?, ?, ?)$

Ex 1 — NOgative example

$\phi$ will not match with big, red, circle. so it is consistent because expected value is -ve. so no change in $S_1$.

$S_1 = (\phi, \phi, \phi)$

General hypothesis ie ? matches with big, red, circle so it is +ve classification and expected is -ve, we need to write all hypothesis which makes it as consistent.

We need to write opposite value of value which is present currently.

Big opposite

$G_1 = (small, ?, ?)(?, Blue, ?)(?, ?, Triangle)$

$S_1 = (\phi, \phi, \phi)$

$G_1 = (small, ?, ?)\ (?, Blue, ?)\ (?, ?, Tri$

Ex - 2 — -ve example

$\phi$ doesn't match with small, Red, Triangle so its -ve classification, expected is also -ve

So $S_2$ remains the same.

ie $S_2 = (\phi, \phi, \phi)$

$G_2$ :

① (Small, ?, ?) — Small matches with small

? " " Red

? " " Triangle

so its +ve classification but expected is -ve

Define possible hypothesis

$G_2$ : (small, Blue, ?) (small, ?, Circle)

opposite to Red — opp to Triangle

② $\phi$ ?, Blue, ?)

— ? matches with small

Blue doesn't match with Red — so -ve classification, expected is also -ve. so we will leave as it.

$G_2$ : (Small, Blue, ?) (small, ?, circle) ( ?, Blue, ?)

$G_2$ - ③ ( ?, ?, Triangle )

? matches with Small

? " " Red

Triangle " " Triangle — +ve classification

but -ve o/p. so possible hypothesis need to find. as it is inconsistent.

B: First ? is replaced with opposite value

The value in ex2 is small, we need to replace opposite value ie Big

( Big, ?, Triangle ) ( ?, Blue, Triangle )

∴ $G_2$: (small, Blue, ?) (small, ?, Circle) ( ?, Blue, ? )
( Big, ?, Triangle ) ( ?, Blue, Triangle )

Among these we need to check which hypothesis matches with the examples what we have seen earlier. ( it should match with both ex1 & ex2 )

$G_2$ - ① (small, Blue, ?)

ex2 { small matches with small
Blue doesn't match with Red } — so -ve classification, expected is -ve
so consistent.

ex1 (small, Blue, ?)
small doesn't match with Big — so -ve, expected is also -ve

llarly go with other hypothesis.

$G_2$ : (small, Blue, ?) (small, ?, circle) (?, Bl

(Big, ?, Triangle) (?, Blue, Trlangle)

Ex-3 – +ve example.

we need to check whether $G_2$ hypothesis is consistent with ex-3, if it is, remain otherwise remove it.

ex3 - (small, Red, Circle)

$G_2$ – ① (small, Blue, ?)

small matches with small
Blue doesn't " with Red
} -ve classification but expected +ve

so inconsistent. Remove

② (small, ?, circle) matches with ex3

so consistent. Remains same

③ (?, Blue, ?) ? matches with small
Blue doesn't match " Red
} - ve classifier but expected +ve

so inconsistent. Remove it

④ (Big, ?, Trlangle) – Big doesn't match with small
? matches with Red
Trlangle doesn't match " circle
} -ve classification, but expected +ve

so inconsistent. Remove it.

?, Blue, Triangle)

? matches with small
Blue doesn't match " Red
Triangle " " " Circle

} -ve classification, but expected is +ve

So inconsistent. Remove it.

So, Consistent hypothesis ~~are~~ is

$G_3$ = (small, ?, circle)

$S_3$ = (small, Red, Circle)

Ex 4 :--ve example Ex 4 (Big, blue, circle)

$S_3 \rightarrow$ Ex 4

~~small~~ doesn't match big
Red " " blue
circle match circle

} -ve classification expected also -ve

So remains same as $S_3$

∴ $S_4$ = (small, red, circle)

Big - small → no match, so retains as it is

So $G_4 = G_3$ = (small, ?, circle)

Ex 5: (small, blue, circle) – +ve example

$G_4$ – (small, ?, circle)

Small matches with small
? " " blue
Circle " " circle

} +ve classi
result is a
+ve

∴ $G_5$ = (small, ?, circle)

$S_5$ = $S_4$ compare with Ex 5

= (small, ?, circle

$S_5$ = (small, ?, circle)

$G_5$ = (small, ?, circle)

# Candidate Elimination Algorithm

| Attendance | Assignment | Exam Prep | Pass |
|---|---|---|---|
| High | Submitted | Good | yes |
| High | Submitted | Poor | yes |
| Low | NotSubmitted | Poor | No |
| High | NotSubmitted | Good | yes |

$S_0 = \langle \emptyset, \emptyset, \emptyset \rangle$

$G_0 = \langle ?, ?, ? \rangle$

$S_1$ = (High, Submitted, Good)
(+ve)

$G_1$ = (?, ?, ?)

$S_2$ = (High, submitted, ?)
(+ve)

$G_2$ = (?, ?, ?)

-ve example

$G_3$ = (<High, ?, ?> <?, Submitted, ?>)

$S_3 = S_2$

$S_3$ = (High, Submitted, ?)

$S_4$ = (High, ?, ?) — $S_4$ - +ve example

$G_4$ = (High, ?, ?)

# Candidate Elimination Algorithm

| Outlook | Temperature | Wind | Picnic |
|---|---|---|---|
| Sunny | Mild | Weak | Yes |
| Cloud | Mild | Weak | Yes |
| Rainy | Cold | Strong | No |
| Sunny | Hot | Weak | Yes |

$S_0 = (\phi, \phi, \phi)$

$G_0 = (? ? ?)$

Ex 1 — +ve example

$S_1 = (Sunny, Mild, Weak)$

$G_1 = (? ? ?)$

Ex-2 : +ve example

$S_2 = (?, Mild, Weak)$

$G_2 = (? ? ?)$

Ex 3 — -ve example

$S_3 = (?, Mild, Weak)$

$G_3 \Rightarrow$ All ? matches with values. So +ve classification but -ve o/p. So inconsistent. We need to find possible hypothesis by replacing with opposite values.

$G_3 = (Sunny, ? ?)(Cloud, ?, ?)(? Mild, ?)(? Hot ?)$
$(? ? Weak)$

(?, Mild, weak)

(sunny, ?, ?) (cloud, ? ?) (? Mild ?)

(? Hot ?) (? ? weak)

Ex 4 — +ve example

$S_4$ = (?, ?, weak)

$G_4$ = {(sunny, ?, ?) (?, Hot, ?) (? ? weak)}

∴ $S_4$ = (?, ? weak)

$G_4$ = {(sunny, ?, ?) (?, Hot, ?) (? ? weak)}

$G_4$ ⇒ remove sunny & Hot

∴ $G_4$ = (?, ?, weak)

# Candidate Elimination Algorithm

| Internet | Devices | Teacher Trained | Adopt |
|---|---|---|---|
| Good | Enough | Yes | Yes |
| Good | Few | Yes | Yes |
| Poor | Few | No | NO |
| Good | Enough | No | Yes |

$S_0 = (\phi, \phi, \phi)$

$G_0 = (? ? ?)$

Ex 1 — +ve example

$S_1 = (Good, Enough, Yes)$

$G_1 = (? ? ?)$

Ex 2 — +ve example

$S_2 = (Good, ?, Yes)$

$G_2 = (? ? ?)$

Ex 3 — -ve example

As all ? matches with values. So +ve Classification, result is -ve, so it is inconsistent. Find all possible hypothesis by replacing with opposite values

$G_3$ = ( Good, ? ?) ( ? Enough ? ) ( ? ? Yes )

$S_3$ = ( Good, ?, Yes )

Ex 4 — +ve

$S_4$ = ( Good, ?, ? )

$G_4$ = { (Good, ? ,?) ( ?, enough, ? ) }

$S_0$ = ( ∅, ∅ ∅ )

$G_0$ = ( ? ? ? )

Ex 1 — -ve.

$S_1$ = ( ∅ ∅ ∅ )

$G_1$ = ( small, ? ?) (Red) ×

$G_1$ = ( small, ? ? )

Ex 2 — -ve $S_2$ = ( ∅ ∅ ∅ )

$G_2$ = ( ( ? ? Circle )

Ex 3 — +ve

$S_3$ = small, red, circle

$G_3$ = ( ? ? Circle )

Ex 4 — -ve

$S_4$ = small, red, circle

$G_4$ = ( small, ? ?) ( ?, re

Ex 5 +ve

$S_5$ = (small, ?, circle)

$G_5$ = (small, ?, ?